

Fabrication of quantum dots in undoped Si/Si_{0.8}Ge_{0.2} heterostructures using a single metal-gate layer

T. M. Lu,^{1, a)} J. K. Gamble,¹ R. P. Muller,¹ E. Nielsen,¹ D. Bethke,¹ G. A. Ten Eyck,¹ T. Pluym,¹ J. R. Wendt,¹ J. Dominguez,¹ M. P. Lilly,¹ M. S. Carroll,¹ and M. C. Wanke¹
Sandia National Laboratories, Albuquerque, New Mexico 87185,
USA

(Dated: 30 August 2016)

Enhancement-mode Si/SiGe electron quantum dots have been pursued extensively by many groups for their potential in quantum computing. Most of the reported dot designs utilize multiple metal-gate layers and use Si/SiGe heterostructures with Ge concentration close to 30%. Here we report the fabrication and low-temperature characterization of quantum dots in Si/Si_{0.8}Ge_{0.2} heterostructures using only one metal-gate layer. We find that the threshold voltage of a channel narrower than 1 μm increases as the width decreases. The higher threshold can be attributed to the combination of quantum confinement and disorder. We also find that the lower Ge ratio used here leads to a narrower operational gate bias range. The higher threshold combined with the limited gate bias range constrains the device design of lithographic quantum dots. We incorporate such considerations in our device design and demonstrate a quantum dot that can be tuned from a single dot to a double dot. The device uses only a single metal-gate layer, greatly simplifying device design and fabrication.

^{a)}Electronic mail: tlu@sandia.gov

Electron spins are one of the most promising candidates for implementing solid-state qubits¹. Si, in particular, is a widely pursued host material for spin qubits, due to its long coherence times² resulting from its weak spin-orbit interaction and the possibility of obtaining isotopically pure ^{28}Si which has zero nuclear spin³. In Si, manipulation of individual electron spins can be achieved in electrostatically defined quantum dots in addition to donor-bound electrons as proposed by Kane⁴. A typical starting platform for making quantum dots is a two-dimensional electron gas (2DEG), which either resides at the Si/SiO₂ interface in the case of Si metal-oxide-semiconductor field-effect transistors (MOSFETs), or in a strained Si quantum well in the case of Si/SiGe heterostructures. A Si/SiGe heterostructure has epitaxial interfaces, which are much less disordered than amorphous SiO₂ on Si and result in much higher electron mobilities. For Si/SiGe heterostructures, one can either use modulation-doping⁵ and fabricate gates to locally deplete electrons, or employ a MOSFET-like enhancement-mode architecture⁶ and induce electrons only where desired. Due to higher starting mobility and better device stability, the enhancement-mode architecture has gained popularity in recent years. Device designs, dot operation, and qubit manipulation have all been reported using this enhancement-mode architecture^{7–10}.

In most of the reported dot designs, the starting Si/SiGe heterostructure has a Ge concentration close to 30%, and the gates used to define quantum dots are structured in multiple metal layers. There has not been much discussion on the choice of this combination. Since electron mobility is the most obvious and convenient gauge for material quality and the record electron mobility¹¹ was demonstrated in a heterostructure with a Ge content of 20%, we chose the same composition for the relaxed SiGe buffer layers as we started fabricating Si/SiGe quantum dots. The higher achievable mobility was expected to reduce the impact of disorder. At first glance, the main difference between Si/Si_{0.7}Ge_{0.3} and Si/Si_{0.8}Ge_{0.2} is simply the barrier height in the 2DEG normal direction. This difference in barrier height is naively not expected to mandate a change in quantum dot design, since most design considerations center around creating lateral confinement of electrons to form quantum dots. Surprisingly, we learned that a reduced Ge concentration affects device performance indirectly but strongly. In this work we study the physical mechanisms of this effect through low-temperature experiments and simulations, and discuss the constraints imposed on device design. Accounting for the constraints, we demonstrate electrical control of quantum dots in a Si/Si_{0.8}Ge_{0.2} heterostructure. Furthermore, the fabrication process flow presented here uses

only a single metal-gate layer, which greatly simplifies device fabrication, shortens device turnaround times, and improves yield by avoiding potentially leaky metal-insulator-metal stacks.

In this work we used undoped Si/Si_{0.8}Ge_{0.2} heterostructures grown by Lawrence Semiconductor Research Laboratory. A virtual substrate was made by growing a 2- μm linearly graded buffer layer on a Si wafer at a grading rate of 10%/ μm , followed by a 1- μm Si_{0.8}Ge_{0.2} relaxed buffer layer. Chemical mechanical polishing (CMP) removed 150 nm of the relaxed buffer layer to reduce surface roughness. After CMP, epi-layers of 400 nm Si_{0.8}Ge_{0.2}, 20 nm Si, 35 nm Si_{0.8}Ge_{0.2}, and 3 nm Si were grown in order. The heterostructures were processed in the Si foundry at Sandia National Laboratories. A similar process flow for SiGe heterostructures has been reported earlier⁷. Electron-beam lithography (EBL) defined the nanostructured gate patterns. 15 nm Al₂O₃ and 100 nm Al were then deposited in the same run, followed by lift-off. The Al₂O₃ layer that insulates the gate from the substrate was formed by depositing Al in the presence of O₂.

We made a series of narrow transistors with variable channel widths, ranging from 100 nm to 5 μm wide. Each narrow gate was 2 μm long. At each end of a narrow gated region, the enhancement gate was flared quickly to a width exceeding 10 μm and was extended to overlap two high doped regions on each end of the narrow wire to make two ohmic contacts on each end. The device threshold voltages were characterized at 4 K by monitoring the onset of channel current at a quasi-dc drain-source bias (V_{DS}) of 1 mV against the gate voltage (V_G). Using the four ohmic contacts for each device, we extracted two threshold voltages, one for the narrow channel, and one for the wide channel. The difference in the two threshold voltages for each device was then tabulated. Some narrow channels did not turn on below $V_G = 2$ V, the highest voltage applied. These devices were excluded from the calculation of threshold shifts. Fig. 1 (c) shows the average shift and the standard deviation of threshold voltages of the narrow channels, and Fig. 1 (d) shows the fraction of working devices. The non-uniform standard deviation arises from outliers that had a threshold voltage much higher than the other devices in the same group.

It is clear from Fig. 1 (c) and (d) that the threshold voltage of a narrow channel shifts higher and a larger fraction of devices do not turn on as the channel gets narrower. In the case of 100-nm-wide channels, none of the devices showed turn-on below $V_G = 2$ V. This is in stark contrast to 100-nm channel widths routinely achieved using similar structures with

a 30% Ge concentration^{9,10}. We attribute the non-working channels and the threshold shifts to the lower Ge concentration used in this work in combination with two other mechanisms, lateral quantum confinement and disorder.

To understand the lateral quantum confinement effect, we performed simulations to analyze the nanowire geometry explored by the experiment. For each wire width, we computed the electrostatic landscape by solving Poisson's equation in 2D with the finite element method in COMSOL Multiphysics using the geometry shown in Fig. 1 (a). We then used this potential energy landscape as input to Schrödinger's equation, and solved for the energies of the 1D nanowire sub-bands. Since different wire widths produce different 2D ground state energies for the same applied voltage, the lowest sub-band begins filling at different applied voltages for different wire widths.

To check if this model is consistent with our experiment, we note that the ground state energy of the 2D system is linear with gate voltage, as shown by the computational results in Fig. 1 (b). Hence, we may write

$$E_0^L = m_L V_G, \quad (1)$$

where E_0^L is the ground state energy of the wire of width L , m_L is the slope (determined computationally), and V_G is the applied gate voltage. By inverting this relationship, we can express the voltage shift of a narrow wire compared to an infinitely wide wire as,

$$\Delta V_L = \frac{E_0}{m_L m_\infty} (m_\infty - m_L), \quad (2)$$

where m_∞ is the slope in the infinite wire-width limit, and E_0 represents a constant threshold energy independent of wire width. We assume the threshold occurs when the ground state energy level of a specific wire, E_0^L , reaches E_0 . Noting that the values of m_L saturate for the wide wires we considered, we take $m_\infty \approx m_{5000} = -930$ meV/V. Treating E_0 as a free parameter and fitting the model to the data using a chi-squared test statistic, we obtained a goodness of fit $p = 0.9953$, indicating excellent consistency with experiment.

In addition to quantum confinement, we believe that disorder also plays an important role in the observed threshold voltage shift. In the presence of disorder, a percolation conduction threshold is associated with each narrow channel. The local conduction threshold may be higher for narrow channels if the widths of the channels are smaller than the characteristic length scale of the disorder potential. Statistical variation mandates that some devices would

experience stronger disorder than others and show higher threshold voltages. The spread of the experimentally observed threshold voltage shifts for a fixed channel width give support to the effect.

The two mechanisms discussed above lead to a higher average threshold voltage as well as a wider spread. This is true for all Si-based narrow channels. For Si MOSFETs, a shift in threshold voltage is not a critical problem. In principle, one can always increase the gate voltage to induce more electrons, and eventually the conduction threshold is overcome at high enough densities. *Unfortunately for Si/SiGe heterostructure FETs, there exists an upper density limit*¹². For shallow Si quantum well channels this maximum density signals the onset of population of a surface channel at the oxide/semiconductor interface. For deep Si quantum wells, a non-equilibrium distribution of electrons collect in the well, with a maximum density controlled by tunneling to a parallel surface channel. A detailed discussion of this non-equilibrium charge distribution can be found in Ref. 12. In both cases, the saturation density is dependent on the barrier height, or in turn, the Ge concentration of the SiGe layers, with the maximum density independent of the channel depth in the latter case¹². The wells used here are considered deep wells and as such, the electron density can be higher than the thermal equilibrium value and is dynamically limited by slow tunneling. Experimentally this upper density limit is $\sim 3 \times 10^{11} \text{ cm}^{-2}$ for $\text{Si}_{0.8}\text{Ge}_{0.2}$ barriers and $\sim 7 \times 10^{11} \text{ cm}^{-2}$ for $\text{Si}_{0.7}\text{Ge}_{0.3}$ barriers¹². We obtained a similar upper density limit in our structures using Hall bar testers fabricated from the same material. This density limit, nonlinear in barrier height, limits the tunable density range in $\text{Si}/\text{Si}_{0.8}\text{Ge}_{0.2}$ to less than half the density range in $\text{Si}/\text{Si}_{0.7}\text{Ge}_{0.3}$. The maximum density results in a maximum Fermi energy. For gate voltages beyond the density saturation bias, a possibly more serious effect occurs, namely, the oxide/semiconductor interface forms a second accumulation layer and eventually accumulates a high enough electron density to become conducting¹³. When the surface channel conduction occurs, this surface channel screens the electric fields to the buried Si quantum well channels, and reduces the electron density in the Si quantum well to its thermal equilibrium value. The thermal equilibrium value could be below the conduction threshold, such that the channel conductivity drops to zero. This appears to happen in all our samples, in which we observe a sudden drop in current to zero at a gate bias beyond threshold, as shown in the inset of Fig. 1 (d).

Since the current in our devices has to flow serially through both the wide and narrow

channel areas, the maximal Fermi energy and the sudden turnoff pose a serious problem if the maximum Fermi energy in the wide gated channel area is not much higher or even lower than the conduction threshold energy of the narrow channel. In the former case, the narrow channel barely turns on before the wide channel turns off, and hence the resistance is very high. In the latter case, when the lateral confinement lifts the threshold energy level significantly, the narrow wire will not reach threshold until the gate is biased beyond the turn-off voltage of the wide channel area. Thus for the narrowest wires there will never be a simultaneous conduction path through both regions. Using Si/Si_{0.8}Ge_{0.2} heterostructures instead of Si/Si_{0.7}Ge_{0.3} thus severely constrains the operational window of V_G . A multiple-metal-layer architecture with independent controls for the reservoirs and nanostructures circumvents this density limit problem since the potential landscape can be locally tuned.

Having understood the major difference between Si/Si_{0.7}Ge_{0.3} and Si/Si_{0.8}Ge_{0.2}, we now turn to our results of fabricating quantum dots using Si/Si_{0.8}Ge_{0.2} heterostructures. Since the main difficulty in using lower Ge concentration is the high threshold voltages for narrow channels, the quantum dot designs have to be enlarged compared to what are used in Si MOSFETs or in Si/Si_{0.7}Ge_{0.3}, unless a multiple-metal-layer architecture is adopted. For this work we chose to use only one metal layer and therefore chose to make large quantum dots. A larger quantum dot has lower charging energies and requires lower temperatures to resolve transport features. However, the larger feature sizes are more easily fabricated. Furthermore, the one-metal-layer architecture has fewer fabrication steps and fewer potentially leaky metal-insulator-metal stacks. This simpler process flow, together with the relaxed nano-patterning accuracy, significantly shortens fabrication turnaround times and improves device yield.

The starting wafer for the dot work had nominally identical growth parameters as previously described. Here we did every step at die-level. Ti/Au alignment marks were first deposited, followed by ion implantation of P at 20 KeV and 75 KeV with a fluence of $5 \times 10^{14} \text{ cm}^{-2}$ for each energy. A rapid thermal anneal at 625 °C for 30 sec activated the implanted dopants for ohmic contacts. A blanket 20-nm-thick Al₂O₃ layer was deposited in an atomic-layer-deposition system at 200 °C. After etching vias through Al₂O₃, we deposited a blanket 2 nm Ti and 40 nm Au film. This metal layer was patterned by EBL and etched in an ion mill. The exposed metal areas were milled away, leaving nanostructure gates and bond pads simultaneously. Fig. 2 (a) shows a scanning electron micrograph of a fabricated quantum-

dot device. Also shown is the circuit setup used for the data presented below. The names of the gates are labeled in yellow. The narrowest point of the channel had a width of ~ 200 nm.

The device shown in Fig. 2 (a) was studied in a ^3He cryostat with a base temperature of ~ 0.3 K at $V_{DS} = 10$ μV using standard lock-in techniques. We focused on the upper channel and varied the voltages for UL, UC, and UR to form quantum dots. The voltages used for AGU, AGL, LL, LC, and LR were kept constant at 0.633, -0.35, 0, 0, and 0 V, respectively, for the data presented here. In Fig. 2 (b) we show a series of stability plots against the UL and UR gates at different UC voltages. At low (less negative) UC, single-quantum-dot behavior is dominant in the lower left corner of a UR vs. UL plot, with approximately equal couplings to the UL and UR gates. Upon making UC more negative, the dot is broken up into two halves by the electrostatic potential, and forms a pair of quantum dots coupled by tunneling. The isolated transport peaks observed showed the expected triple points for tunnel-coupled double-quantum-dots¹⁴. From stability diagrams, we extract the following gate-dot capacitances for the double-dot system: $C_{UL,L} = 4.0$ aF, $C_{UC,L} = 5.2$ aF, $C_{UR,L} = 1.2$ aF, $C_{AGU,L} = 63$ aF, $C_{AGL,L} = 13$ aF, $C_{LL,L} = 1.3$ aF, $C_{LC,L} = 1.4$ aF, $C_{LR,L} = 0.48$ aF, $C_{UL,R} = 1.3$ aF, $C_{UC,R} = 5.2$ aF, $C_{UR,R} = 4.0$ aF, $C_{AGU,R} = 79$ aF, $C_{AGL,R} = 13$ aF, $C_{LL,R} = 0.63$ aF, $C_{LC,R} = 1.3$ aF, $C_{LR,R} = 0.48$ aF. These capacitances are consistent with a lithographic double-dot defined by the gates. The inter-dot tunnel barrier is not only tuned by the UC gate, but also by the UL and UR gates, as is evidenced from the transition from single-dot-like to double-dot-like characteristics moving from the upper right corner to the lower left corner at a fixed UC voltage. This is a trade-off of using a single-metal-gate-layer design; controls over tunnel barriers and dot occupations are more intertwined together and require more sophisticated device tuning.

To corroborate our single to double-dot interpretation of Fig. 2 (b), we performed capacitive modeling of our nanostructure using FastCap to obtain a simulated charge stability diagram. We used voltages identical to our experiment, modeled the electrical leads below AGU as a metallic brick, and the two quantum dots as 10 nm thick conductive bricks with 400×400 nm in-plane dimensions, as shown in Fig. 2 (a). The two dots were separated by a distance d , which we varied. As shown in Fig. 2 (c), by changing the distance from 0.1 to 10 nm, we are able to achieve a clear transition from a single- to double-dot charge stability signature. These plots are qualitatively consistent with those in Fig. 2 (b), indicating that

one needs only to consider (plausible) small dot separations when explaining the data.

In summary, we discovered the constraints imposed by using Si/SiGe heterostructures with a Ge concentration of 20%, lower than what is used in other reports. We found that the limited tunable density range prevents conduction through channels gated by a single layer gate which contains both very narrow and wide channel widths. With a scaled-up quantum dot design, we demonstrated operation of lithographically defined quantum dots using Si/Si_{0.8}Ge_{0.2} heterostructures, and were able to induce a single-dot to double-dot transition. The presented fabrication process flow requires only one metal-gate-layer for all gates and bond pads, significantly reducing device fabrication turnaround times. This may find use in situations where high throughput of simple quantum dots is required, such as studying the statistics of quantum dot properties.

This work was performed, in part, at the Center for Integrated Nanotechnologies, a U.S. DOE, Office of Basic Energy Sciences, user facility. Sandia National Laboratories is a multi program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. DOE's National Nuclear Security Administration under contract DE-AC04-94AL85000.

REFERENCES

- ¹D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).
- ²A. M. Tyryshkin, S. A. Lyon, A. V. Astashkin, and A. M. Raitsimring, Phys. Rev. B **68**, 193207 (2003).
- ³F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, Rev. Mod. Phys. **85**, 961 (2013).
- ⁴B. E. Kane, Nature **393**, 133 (1998).
- ⁵F. Schffler, Semicond. Sci. Tech. **12**, 1515 (1997).
- ⁶T. M. Lu, J. Liu, J. Kim, K. Lai, D. C. Tsui, and Y. H. Xie, Applied Physics Letters **90**, 182114 (2007).
- ⁷T. M. Lu, N. C. Bishop, T. Pluym, J. Means, P. G. Kotula, J. Cederberg, L. A. Tracy, J. Dominguez, M. P. Lilly, and M. S. Carroll, Appl. Phys. Lett. **99**, 043101 (2011).
- ⁸D. Kim, Z. Shi, C. Simmons, D. Ward, J. Prance, T. S. Koh, J. K. Gamble, D. Savage,

- M. Lagally, M. Friesen, S. N. Coppersmith, and M. A. Eriksson, *Nature* **511**, 70 (2014).
- ⁹K. Eng, T. D. Ladd, A. Smith, M. G. Borselli, A. A. Kiselev, B. H. Fong, K. S. Holabird, T. M. Hazard, B. Huang, P. W. Deelman, I. Milosavljevic, A. E. Schmitz, R. S. Ross, M. F. Gyure, and A. T. Hunter, *Science Advances* **1**, e1500214 (2015).
- ¹⁰D. M. Zajac, T. M. Hazard, X. Mi, K. Wang, and J. R. Petta, *Appl. Phys. Lett.* **106**, 223507 (2015).
- ¹¹T. M. Lu, D. C. Tsui, C.-H. Lee, and C. W. Liu, *Appl. Phys. Lett.* **94**, 182102 (2009).
- ¹²T. M. Lu, C.-H. Lee, S.-H. Huang, D. C. Tsui, and C. W. Liu, *Appl. Phys. Lett.* **99**, 153510 (2011).
- ¹³C.-T. Huang, J.-Y. Li, K. S. Chou, and J. C. Sturm, *Appl. Phys. Lett.* **104**, 243510 (2014).
- ¹⁴W. G. van der Wiel, S. De Franceschi, J. M. Elzerman, T. Fujisawa, S. Tarucha, and L. P. Kouwenhoven, *Rev. Mod. Phys.* **75**, 1 (2002).

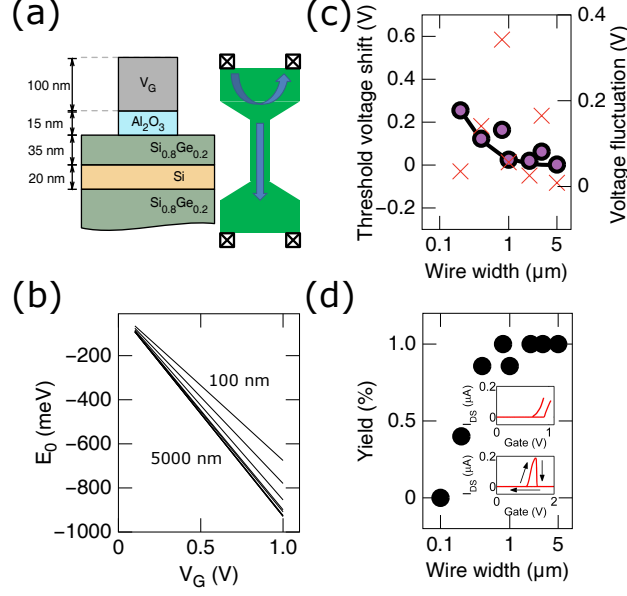


FIG. 1. (a): Schematic cross-section and top view of the measured/modelled FET. The arrows illustrate the current paths of the wide and narrow channels. (b): Calculated ground state energy of the channel produced by applying a voltage on the gate indicated in (a). The lines correspond to the wire widths measured in panel (c). (c): Threshold voltage shift (solid dots) and its standard deviation (x markers) as a function of wire width. The line is the model fitted to a single-parameter model, obtaining a goodness of fit $p = 0.9953$. (d): Fraction of functional devices (yield) as a function of wire width. Top inset: representative turn-on curves at $V_{DS} = 1$ mV for a 200-nm wide channel (right curve) and an 800-nm wide channel (left curve). Bottom inset: a representative turn-on curve showing that the channel shuts off beyond a critical gate voltage.

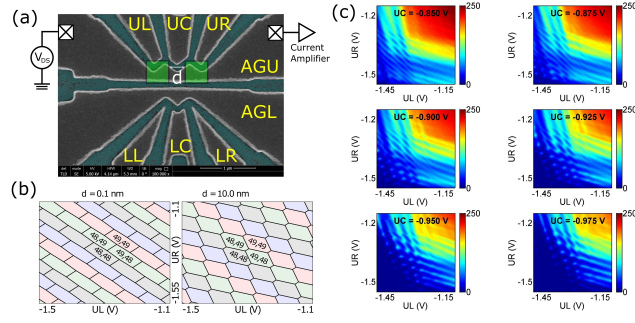


FIG. 2. (a) Scanning electron micrograph of the device. Lighter gray with yellow labels are Ti/Au gates on top of the Si/SiGe substrate. The cyan pseudo-colored regions are exposed Al_2O_3 . The capacitance calculations assume two square quantum dots with side length 400 nm as indicated in green. These two dots are separated by a distance d . (b) Stability diagrams for the quantum dot in the upper channel at $V_{UC} = -0.85, -0.875, -0.9, -0.925, -0.95, \text{ and } -0.975$ V. The measurement temperature was 0.3 K, and the drain-source bias was 10 μV . (c) Capacitance simulations for differing dot separation d . We see that even a modest dot separation of 10 nm is sufficient to transition from single- to double-dot behavior. These calculations indicate that the observed charge stability diagrams in (b) likely describe a transition between lithographic dots.